Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u> <u>SUGGESTED CORRECTION</u>   SERIAL NUMBER: _O9/424,487A_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics   The number/text at the end of each line "wrapped" down to the next line.
                               This may occur if your file was retrieved in a word processor after creating it.
                               Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos   The amino acid number/text at the end of each line "wrapped " down to the next line.
                               This may occur if your file was retrieved in a word processor after creating it.
                               Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
        Numbering   between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII   This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
                               Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length   Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
                               As per the rules, each n or Xaa can only represent a single residue.
                               Please present the maximum number of each residue having variable length and
                               indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
                               sequence(s) _____. Normally, PatentIn would automatically generate this section from the
                               previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
                               to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
                               sections for Artificial or Unknown sequences.

8 _____ Skipped Sequences   Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
        (OLD RULES)   (2) INFORMATION FOR SEQ ID NO:X:
                               (i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
                               (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
                               This sequence is intentionally skipped

                               Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences   Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
        (NEW RULES)   <210> sequence id number
                               ‹400> sequence id number
                               000

10 _____ Use of n's or Xaa's   Use of n's and/or Xaa's have been detected in the Sequence Listing.
         (NEW RULES)   Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
                               In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism Sequence(s) _____ are missing this mandatory field or its response.
         (NEW RULES)   4,7-9
12 _____ Use of <220>Feature   Sequence(s) _1_ are missing the <220>Feature and associated headings.
         (NEW RULES)   Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
                               Please explain source of genetic material in <220> to <223> section.
                               (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
                               file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
                               Instead, please use "File Manager" or any other means to copy file to floppy disk.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/424,487A

DATE: 03/28/2001
TIME: 15:24:58

Input Set : N:\Crf3\Refhold\I424487A.raw
Output Set: N:\CRF3\03282001\I424487A.raw

Does Not Comply
Corrected Diskette Needed

```
 1 <110> APPLICANT: CHOO, Yen
 2        KLUG, Aaron
 3        ISALAN, Mark
 4 <120> TITLE OF INVENTION: Nucleic Acid Binding Proteins
 5 <130> FILE REFERENCE: 71278/264975
```
C--> ` 6 <140> CURRENT APPLICATION NUMBER: US/09/424,487A`
```
 7 <141> CURRENT FILING DATE: 2000-02-29
 8 <150> PRIOR APPLICATION NUMBER: GB 9710809.6
 9 <151> PRIOR FILING DATE: 1997-05-23
10 <150> PRIOR APPLICATION NUMBER: PCT/GB98/01512
11 <151> PRIOR FILING DATE: 1998-05-26
12 <160> NUMBER OF SEQ ID NOS: 17
13 <170> SOFTWARE: PatentIn Ver. 2.1
15 <210> SEQ ID NO: 1
16 <211> LENGTH: 264
17 <212> TYPE: DNA
18 <213> ORGANISM: Artificial Sequence
19 <220> FEATURE:
20 <221> NAME/KEY: CDS
21 <222> LOCATION: (1)..(264)
22 <223> OTHER INFORMATION: Description of Artificial Sequence: encoding
23        nucleic acid binding proteins
24 <400> SEQUENCE: 1
25       gca gaa gag aag cct ttt cag tgt cga atc tgc atg cgt aac ttc agc    48
26       Ala Glu Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser
27        1               5                  10                  15
28       gat cgt act act ctt acc cgc cac acg agg acc cac aca ggc gag aag    96
29       Asp Arg Thr Thr Leu Thr Arg His Thr Arg Thr His Thr Gly Glu Lys
30                    20                  25                  30
31       cct ttt cag tgt cga atc tgc atg cgt aac ttc agc agg agc gat aac   144
32       Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn
33                35                  40                  45
34       ctt acg aga cac cta agg acc cac aca ggc gag aag cct ttt cag tgt   192
35       Leu Thr Arg His Leu Arg Thr His Thr Gly Glu Lys Pro Phe Gln Cys
36            50                  55                  60
37       cga atc tgc atg cgt aac ttc agg caa gct gat cat ctt caa gag cac   240
38       Arg Ile Cys Met Arg Asn Phe Arg Gln Ala Asp His Leu Gln Glu His
39        65                  70                  75                  80
40       cta aag acc cac aca ggc gag aag                                   264
41       Leu Lys Thr His Thr Gly Glu Lys
42                    85
44 <210> SEQ ID NO: 2
45 <211> LENGTH: 88
46 <212> TYPE: PRT
47 <213> ORGANISM: Artificial Sequence
48 <220> FEATURE:
49 <223> OTHER INFORMATION: Description of Artificial Sequence: encoding
```

```
 50        nucleic acid binding proteins
 51 <400> SEQUENCE: 2
 52        Ala Glu Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser
 53         1               5                   10                  15
 54        Asp Arg Thr Thr Leu Thr Arg His Thr Arg Thr His Thr Gly Glu Lys
 55                        20                  25                  30
 56        Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn
 57                        35                  40                  45
 58        Leu Thr Arg His Leu Arg Thr His Thr Gly Glu Lys Pro Phe Gln Cys
 59              50                  55                  60
 60        Arg Ile Cys Met Arg Asn Phe Arg Gln Ala Asp His Leu Gln Glu His
 61         65                  70                  75                  80
 62        Leu Lys Thr His Thr Gly Glu Lys
 63                        85
 65 <210> SEQ ID NO: 3
 66 <211> LENGTH: 18
 67 <212> TYPE: PRT
 68 <213> ORGANISM: Artificial Sequence
 69 <220> FEATURE:
 70 <223> OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
 71       binding protein
 72 <221> NAME/KEY: BINDING
 73 <222> LOCATION: (1)..(18)
 74 <223> OTHER INFORMATION: where X is any amino acid
 75 <400> SEQUENCE: 3
 76        Xaa Cys Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Xaa Xaa
 77         1               5                   10                  15
 78        Xaa His
 80 <210> SEQ ID NO: 4
 81 <211> LENGTH: 21
 82 <212> TYPE: PRT
 83 <213> ORGANISM: Artificial Sequence
 84 <220> FEATURE:
 85 <223> OTHER INFORMATION: Description of Artificial Sequence: where X is any
 86       amino acid
 87 <221> NAME/KEY: BINDING
 88 <222> LOCATION: (1)..(21)
 89 <400> SEQUENCE: 4
 90        Xaa Cys Xaa Xaa Cys Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
 91         1               5                   10                  15
 92        His Xaa Xaa Xaa His
 93                        20
 95 <210> SEQ ID NO: 5
 96 <211> LENGTH: 26
 97 <212> TYPE: PRT
 98 <213> ORGANISM: Artificial Sequence
 99 <220> FEATURE:
100 <223> OTHER INFORMATION: Description of Artificial Sequence: Consensus zinc
101       finger structure
```

[Handwritten annotations:] This does not describe the genetic source of Artificial Sequence. See item 12 on Error Summary Sheet.

This goes on separate C2237 line

```
102 <221> NAME/KEY: BINDING
103 <222> LOCATION: (1)..(26)
104 <400> SEQUENCE: 5
105      Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Lys Ser Asp
106       1               5                   10                  15
107      Leu Val Lys His Gln Arg Thr His Thr Gly
108                      20                  25
110 <210> SEQ ID NO: 6
111 <211> LENGTH: 29
112 <212> TYPE: PRT
113 <213> ORGANISM: Artificial Sequence
114 <220> FEATURE:
115 <223> OTHER INFORMATION: Description of Artificial Sequence: Consensus
116       zinc finger structure
117 <221> NAME/KEY: BINDING
118 <222> LOCATION: (1)..(29)
119 <400> SEQUENCE: 6
120      Pro Tyr Lys Cys Ser Glu Cys Gly Lys Ala Phe Ser Gln Lys Ser Asn
121       1               5                   10                  15
122      Leu Thr Arg His Gln Arg Ile His Thr Gly Glu Lys Pro
123                      20                  25
125 <210> SEQ ID NO: 7
126 <211> LENGTH: 9
127 <212> TYPE: PRT
128 <213> ORGANISM: Artificial Sequence
129 <220> FEATURE:
130 <223> OTHER INFORMATION: Description of Artificial Sequence: where x
131       denotes a given combination of the bases at
132       interface between DNA subsites, and the four bases
133       are equally represented at DNA position 3
134 <221> NAME/KEY: BINDING
135 <222> LOCATION: (1)..(9)
136 <400> SEQUENCE: 7
137      Gly Asn Xaa Xaa Cys Gly Gly Cys Gly
138       1               5
140 <210> SEQ ID NO: 8
141 <211> LENGTH: 9
142 <212> TYPE: PRT
143 <213> ORGANISM: Artificial Sequence
144 <220> FEATURE:
145 <223> OTHER INFORMATION: Description of Artificial Sequence: where X
146       denotes a known combination of the two bases at
147       DNA positions 4X and 5X and there is equal
148       probability of any of the four bases at DNA position 3
149 <221> NAME/KEY: BINDING
150 <222> LOCATION: (1)..(9)
151 <400> SEQUENCE: 8
152      Gly Asn Xaa Xaa Cys Gly Gly Cys Gly
153       1               5
```

W--> 137

W--> 152

(handwritten annotations)

1) This statement does not explain "Artificial Sequence;

2) Xaa can only be a single amino acid, not a base.

same error

RAW SEQUENCE LISTING   DATE: 03/28/2001
PATENT APPLICATION: US/09/424,487A   TIME: 15:24:58

Input Set : N:\Crf3\Refhold\I424487A.raw
Output Set: N:\CRF3\03282001\I424487A.raw

```
155 <210> SEQ ID NO: 9
156 <211> LENGTH: 9
157 <212> TYPE: PRT
158 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Description of Artificial Sequence; where X
161       denotes a known combination of the two bases at
162       DNA positions 4X and 5X
163 <221> NAME/KEY: BINDING
164 <222> LOCATION: (1)..(9)
165 <400> SEQUENCE: 9
166       Gly Cys Xaa Xaa Cys Gly Gly Cys Gly
167       1               5
169 <210> SEQ ID NO: 10
170 <211> LENGTH: 28
171 <212> TYPE: PRT
172 <213> ORGANISM: Artificial Sequence
173 <220> FEATURE:
174 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
175       binding protein
176 <221> NAME/KEY: BINDING
177 <222> LOCATION: (1)..(28)
178 <400> SEQUENCE: 10
179       Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Arg Ser Ser Leu
180       1               5                   10                  15
181       Thr Arg His Thr Arg Thr His Thr Gly Glu Lys Pro
182                       20                  25
184 <210> SEQ ID NO: 11
185 <211> LENGTH: 28
186 <212> TYPE: PRT
187 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
190       binding protein
191 <221> NAME/KEY: BINDING
192 <222> LOCATION: (1)..(28)
193 <400> SEQUENCE: 11
194       Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Arg Ser His Leu
195       1               5                   10                  15
196       Thr Arg His Thr Arg Thr His Thr Gly Glu Lys Pro
197                       20                  25
199 <210> SEQ ID NO: 12
200 <211> LENGTH: 27
201 <212> TYPE: PRT
202 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
205       binding protein
206 <221> NAME/KEY: BINDING
```

W--> *(handwritten annotations: "same err" and bracket markings)*

```
207 <222> LOCATION: (1)..(27)
208 <400> SEQUENCE: 12
209     Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Asp Arg Ser Asn Leu
210      1               5                  10                  15
211     Thr Arg His Thr Arg Thr His Thr Gly Glu Lys
212                    20                  25
214 <210> SEQ ID NO: 13
215 <211> LENGTH: 9
216 <212> TYPE: PRT
217 <213> ORGANISM: Artificial Sequence
218 <220> FEATURE:
219 <223> OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
220        binding protein
221 <221> NAME/KEY: BINDING
222 <222> LOCATION: (1)..(9)
223 <400> SEQUENCE: 13
224     Ala Gly Ala Gly Ala Gly Cys Thr Cys
225      1               5
227 <210> SEQ ID NO: 14
228 <211> LENGTH: 8
229 <212> TYPE: PRT
230 <213> ORGANISM: Artificial Sequence
231 <220> FEATURE:
232 <223> OTHER INFORMATION: Description of Artificial Sequence: 8bp
233        palindromic sequence which is bound and cleaved by
234        the restriction endonuclease Notl
235 <221> NAME/KEY: BINDING
236 <222> LOCATION: (1)..(8)
237 <400> SEQUENCE: 14
238     Gly Cys Gly Gly Cys Cys Gly Cys
239      1               5
241 <210> SEQ ID NO: 15
242 <211> LENGTH: 9
243 <212> TYPE: PRT
244 <213> ORGANISM: Artificial Sequence
245 <220> FEATURE:
246 <223> OTHER INFORMATION: Description of Artificial Sequence: zinc finger
247        binding protein
248 <221> NAME/KEY: BINDING
249 <222> LOCATION: (1)..(9)
250 <400> SEQUENCE: 15
251     Gly Cys Gly Gly Cys Cys Gly Cys Gly
252      1               5
254 <210> SEQ ID NO: 16
255 <211> LENGTH: 8
256 <212> TYPE: PRT
257 <213> ORGANISM: Artificial Sequence
258 <220> FEATURE:
259 <223> OTHER INFORMATION: Description of Artificial Sequence: Notl
```

VERIFICATION SUMMARY   DATE: 03/28/2001
PATENT APPLICATION: US/09/424,487A   TIME: 15:24:59

Input Set : **N:\Crf3\Refhold\I424487A.raw**
Output Set: **N:\CRF3\03282001\I424487A.raw**

L:6 M:270 C: Current Application Number differs, Wrong Format
L:76 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:78 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:90 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:92 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:137 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:152 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:166 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9